DOCUMENT RESUME

ED 040 754 PS 003 120

AUTHOR Holland, James G.
TITLE The Misplaced Adaptation to Individual Differences.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
PUB DATE 69
NOTE 6p.; Paper originally presented as a part of the symposium at the 1969 convention of the American Psychological Association, Washington, D.C.

EDRS PRICE EDRS Price MF-$0.25 HC-$0.40
DESCRIPTORS Computer Assisted Instruction, *Educational Objectives, *Individualized Instruction, *Program Evaluation, *Testing Problems, Test Reliability, Test Validity
IDENTIFIERS Individually Prescribed Instruction

ABSTRACT

The current interest in an educational technology that stresses adaptation to individual student differences has resulted in such individualized systems as Individually Prescribed Instruction (IPI) and Computer Assisted Instruction (CAI). However, such systems are not the answer to an avoidance of standardization of students. No one has yet resolved the basic dilemma between the cost in time-and-effort efficiency and the demands of test theory for validity and reliability of tests. This dilemma means that while good teaching items should have a low error factor in order to elicit the correct response and then reinforce it, good diagnostic items (needed for individualization) should not have a low error factor. Thus, good teaching items meet criteria incompatible with those met by good diagnostic items. This problem is not insurmountable, but none of the existing programs of individualized instruction have solved it. The most persuasive point against the existing programs is that they are individualized only in terms of what the student brings to the lesson; they still result in the production of uniformity. (MH)

THE MISPLACED ADAPTATION TO INDIVIDUAL DIFFERENCES[1]

JAMES G. HOLLAND
UNIVERSITY OF PITTSBURGH

Today there is a great deal of interest in an educational technology which stresses adaptation to individual differences. It is feared that programming and the increased processing of masses in the name of education will obliterate all divergence and individuality by making the student a mere replica of the standard student defined in our behavioral objectives. Emphasis on systems, such as Individually Prescribed Instruction or Computer Assisted Instruction, that adapt to differences is often taken as an answer to this problem. I suggest that it is not an answer. The fear of deadly uniformity is as great with so-called individualized instruction as without it. To adjust to individual differences in most IPI and CAI systems, continual small modifications of teaching-sequence are made as the result of repeated measures of individual differences. And, although this procedure is taken as "obviously" worthy, I suggest on the one hand, that there are little compelling data to support its "obvious" merit and on the other, that no one, as yet, has resolved a basic dilemma between the cost in efficiency (in terms of time and effort) and the demands of test theory for validity and reliability of tests.

First consider the question of cost efficiency versus reliability of testing. While the rationale for the design of teaching and testing items in IPI and CAI is not often clearly stated and perhaps not universally agreed upon, it seems that generally the rationale for teaching sequences draws directly on that of programmed instruction.

For example, Suppes (1966) acknowledges that teaching material in CAI is based on the rationale of programmed instruction. To be sure, many programmers are very lax in program design and rely on testing to keep those individuals, whom we may consider to be failures of the program, working on additional material until mastery. Generally there is agreement that the teaching elements are supposed to follow good programming principles. Ideally the experiences given individually are arranged in a gradual progression. In the case of individualization, the effort is to uniquely tailor the particular progression to the individual to accomodate his learning rate or his past experience. Individual teaching items, when functioning as teaching items, are expected to require appropriate behavior before the student reaches a correct answer; thus a low black-out ratio (Holland, 1967) is required. And, if functioning as teaching items, they must be able to elicit the behavior to be learned; thus they must have a low error rate. The teaching item does not trap the student into errors or diagnose his difficulties, rather it evokes the desired behavior so that it may be reinforced and established. But individualization requires another type of item. It requires test items which serve a diagnostic function. Such items serve to differentially predict. They are designed to recommend different things for different individuals. Considerations in test design should be brought to bear on the design of diagnostic material for individualization. First, to be useful as a diagnostic item, an item must discriminate among individuals. They must perform differently on it. Some must answer correctly; others

must be in error. A low error rate item does not differentiate individuals. Thus a good diagnostic item meets criteria incompatible with those met by a good teaching item. There is, then, an incompatibility between teaching items and diagnostic items, or at least there is an incompatibility between teaching and diagnostic functions.

To be useful as a diagnostic test, a test must have validity and reliability. Of special importance to us is the relationship between reliability and test length. In general, the shorter the test, the poorer its reliability and validity.

A test enabling one to make a good decision normally requires a reasonable number of items. This problem is openly recognized in the development of IPI. Lindvall and Cox (1969), writing placement tests in IPI note that they require five to ten items for each behavioral objective. Often, especially in CAI, a heavy predictive burden is placed on a single multiple choice item. Consider an item from the Stanford CAI reading program (Atkinson, 1968):

In an item presented by Atkinson (1968) as a typical one from the Stanford CAI reading program the child is shown the letter "r" to the left of an empty cell, and the letters "an" above the cell, and four alternative words listed below ("rat," "bat," "fan," "ran"). The student's response is to touch one of the words. If the correct word is touched, the next mainline item is presented. If an incorrect alternative is touched, an item is presented which requires either choosing the appropriate final consonant (if "rat" was touched) or the appropriate initial consonant (if "fan" was touched) or each of the corrective items in turn (if "bat" was touched).

After the corrective item or items the original mainline item is repeated.

In its diagnostic function, the mainline item which initiates the branching has severe limitations as a four-alternative item. There is a twenty-five per cent opportunity to be correct by chance, and each alternative provides a different course of action. The reliability of such an item (though nowhere stated) must be low. But the consequence is small also; the student receives one or two items and then repeats the missed item. This brings us to the other component of the dilemma of individualization -- the size of the consequence. The size of the decision so far as the student is concerned may vary greatly. How many items must a student do as a function of failing the diagnostic unit? or how much time must he spend? In traditional testing, consequences are often large; a child is given an intelligence test to determine whether he should be in a school for the retarded, a student takes the college boards to determine whether he will profit from college. With such large decisions, it would be a waste to use even a single diagnostic item -- even one of perfect reliability to spare half the students only a single unnecessary teaching item. Moreover, single diagnostic items are seldom very reliable. Thus, the basic technological problem to be worked out for individualization is one of a type of cost efficiency for the student based on the cost in time of the diagnostic material and the resulting savings in the decisions.

A small decision, as frequently made in the constantly adjusting sequence of a CAI program can not justify a lengthy diagnosis, but the short tests are unreliable and the estimate of cost efficiency

must take into account the false positives and false negatives. The more finely grained the adjustment the worse the dilemma, with single test items being unreliable, and, even if they could be reliable, requiring at least two-item consequence to simply break even.

The present great stress on individualization takes considerably more student time in testing to find the teaching items they don't need, than the items themselves would have taken. In an IPI system, a student may spend as much as two-fifths of his time taking tests. With this much testing, the results might be reliable (although reliability and validity information is seldom provided), but the consequences seem insufficient to justify so much testing (as compared, for example, with a once-a-year placement test).

Individualization , to escape this dilemma will require development of measures of cost effectiveness corrected for empirically determined reliabilities. It seems unlikely that many of the more well-known programs would fare well against opposite criteria of this type.

But there is yet a more serious failure of individualization to live up to its press. Today, adjustment for individuality needs to produce greater diversity, rather than less. Programs branch on past achievement, on learning rate, on latency in answering items, but they do not branch on desired terminal behavior. Educational objectives are set by an educational establishment and create a homogeneous middle class. In Skinner's Walden Two, a visiting critic argues against behavioral design in society by objecting to the uniformity he imagines to be a natural consequence of such designs. He is answered by Walden TWO's founder, Frazier, that creation of such uniformity would be bad behavioral engineering. It seems that there are few Fraziers in

today's world of behavioral technology. And the demands of Blacks, of radical students, and others are often for the pursuit of different objectives. To sell individualization as a panacea for these problems is a sham. Branching should be on the basis of different desired outcomes rather than on different entering behavior.

REFERENCES

Atkinson, R.C. Computerized instruction and the learning process. American Psychologist, 1968, 23, 225-239.

Holland, J.G. A quantitative measure for programmed instruction. American Educational Research Journal, 1967, 4, 87-101.

Lindvall, C.M. and Cox, R.C. The role of evaluation in programs for individualized instruction. In Educational Evaluation: New Roles, New Means. 68th Yearbook, NSSL, Part II. Chicago: University of Chicago Press, 1969. Pp. 156-188.

Skinner, B.F. Walden Two. New York: Macmillan, 1948.

Suppes, P. The uses of computers in education. Scientific American, 1966, 215, 206-221.

FOOTNOTES

[1] This chapter was originally presented as a part of a symposium at the 1969 convention of the American Psychological Association. The research reported herein was supported by the Learning Research and Development Center at the University of Pittsburgh, supported in part as a research and development center by funds from the United States Department of Health, Education, and Welfare.